3A

УДК 004.89:004.93

Д.С. Азаренко, С.Б. Иванова, И.Г. Герасимов
Институт проблем искусственного интеллекта
МОН Украины и НАН Украины, г. Донецк
Украина, 83048, г. Донецк, ул. Артема, 118 б

Голосовой калькулятор

D.S. Azarenko, S.B. Ivanova, I.G. Gerasimov
Institute of Artificial Intelligence MES Ukraine and NAS Ukraine, Donetsk
Ukraine, 83048, Donetsk, st. Artem, 118-b

Voice calculator

Д.С. Азаренко, С.Б. Іванова, І.Г. Герасимов
Інститут проблем штучного інтелекту МОН України і НАН України, м. Донецьк
Україна, 83048, м. Донецьк, вул. Артема, 118-б

Голосовий калькулятор

Разработан метод генерация словаря распознавания в реальном времени в зависимости от предыдущих команд.
Ключевые слова: голосовой калькулятор, математическая форма, голосовая команда, список распознавания.

Developed a method for the generation of recognition vocabulary in real time according to the preceding instructions.
Key words: voice calculator, mathematical form, voice command, recognition list.

Розроблено метод генерація словника розпізнавання в реальному часі залежно від попередніх команд.
Ключові слова: голосовий калькулятор, математична форма, голосова команда, список розпізнавання.

Введение

Программа «Голосовой калькулятор» (ГК) предназначена для голосового ввода и вычисления несложных математических выражений. Она может быть полезна для людей с ограниченными возможностями, а также лицам, которым необходимо производить вычисления, но нет возможности вводить формулы при помощи рук. Существует аналог данной программы для ОС Android «Счетовод» [1], существенным недостатком которой является использование он-лайн сервиса Google для распознавания голосовых команд, для чего необходимо постоянное подключение к интернету. Первоочередной задачей является улучшение качества распознавания команд. Решению этой задачи посвящена данная статья.

Результаты исследования

Программа позволяет производить вычисления математических выражений, введенных при помощи голоса. На рис. 1 представлена общая схема функционирования программы ГК.

Ядро ГК (ЯГК) передает для распознавания поступившую от пользователя «голосовую команду» в «модуль распознавания» вместе со «списком распознавания»

(сформированный список текстовых команд, которые могут быть введены в данный момент). Для распознавания команд используется описанный в статьях [2-4] модуль распознавания голосовых команд, посредствам которого производится их преобразование в текст. Результатом распознавания является текстовая команда, которая в ЯГК преобразуется в математическую форму (при помощи «вспомогательных таблиц»). Полученный результат ЯГК добавляет к ранее введенной информации и предоставляется пользователю на экране.

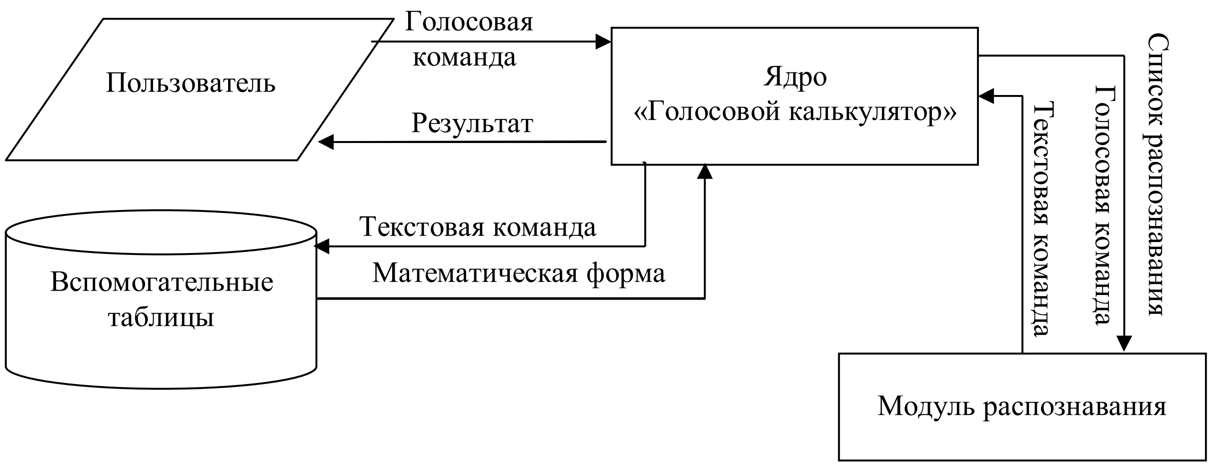


Рисунок 1 – Схема функционирования программы «Голосовой калькулятор»

С целью уменьшения общего числа команд разработан алгоритм автоматического определения, какую скобку (открытую или закрытую) необходимо установить при вводе. Для ввода скобки в программе предусмотрена одна команда – «скобка». Программа автоматически определяет, какую из них необходимо поставить по следующим правилам: в случае начала формулы или после любой математической операции ставится знак «(» (открытая скобка), а после числа ставится знак «)» (закрытая скобка).

Все команды условно разделены на три группы (табл. 1), что обеспечивает возможность сокращения их числа в «списке распознавания».

Сложные числа вводятся последовательно по одной цифре. Например, число «двенадцать» вводиться при помощи последовательности двух команд «один» и «два».

В «списке распознавания» по умолчанию («исходный список распознавания» (ИСР)) присутствуют «числовые команды», «управляющие команды» и команда «скобка», так как эти команды необходимы на любом этапе работы, «Инженерные команды» добавляются в «список распознавания» только после ввода соответствующей команды.

Количество всех команд используемых ГК составляет 40. Для уменьшения количества текстовых команд в «списке распознавания» разработан алгоритм, при помощи которого после ввода команды происходит его формирование на основании ранее введенных данных. При этом добавляются только те команды, которые могут следовать далее (например, после математической команды не может следовать еще одна математическая команда, после команды «плюс» не может следовать «минус»).

Таблица 1 – Группы команд «ГК»

Группа	«Команда» (математическая форма)	Комментарий
«числовые команды»	«ноль» (0) «один» (1) «два» (2) «три» (3) «четыре» (4) «пять» (5) «шесть» (6) «семь» (7) «восемь» (8) «девять» (9) «точка» (.)	команды ввода чисел
«математические команды» арифметические инженерные	«плюс» (+) «минус» (–) «умножить» (*) «разделить» (/) «равно» (=) «скобка» ((или)) «квадрат» (x^2) «куб» (x^3) «степень» (x^y) «корень» ($\sqrt{x}$) «кубический корень» $\sqrt[3]{x}$ «факториал» (!) «синус» (sin) «арксинус» (arcsin) «косинус» (cos) «арккосинус» (arccos) «тангенс» (tan) «арктангенс» (arctan) «логарифм» (lg)	команды математических операций
«управляющие команды»	«инженерный» «калькулятор» «отменить» «обучить» «сбросить» «удалить»	команды, при помощи которых производится обучение модуля распознавания и совершается навигация по веденной формуле

На рис. 2 представлен алгоритм формирования «списка распознавания».

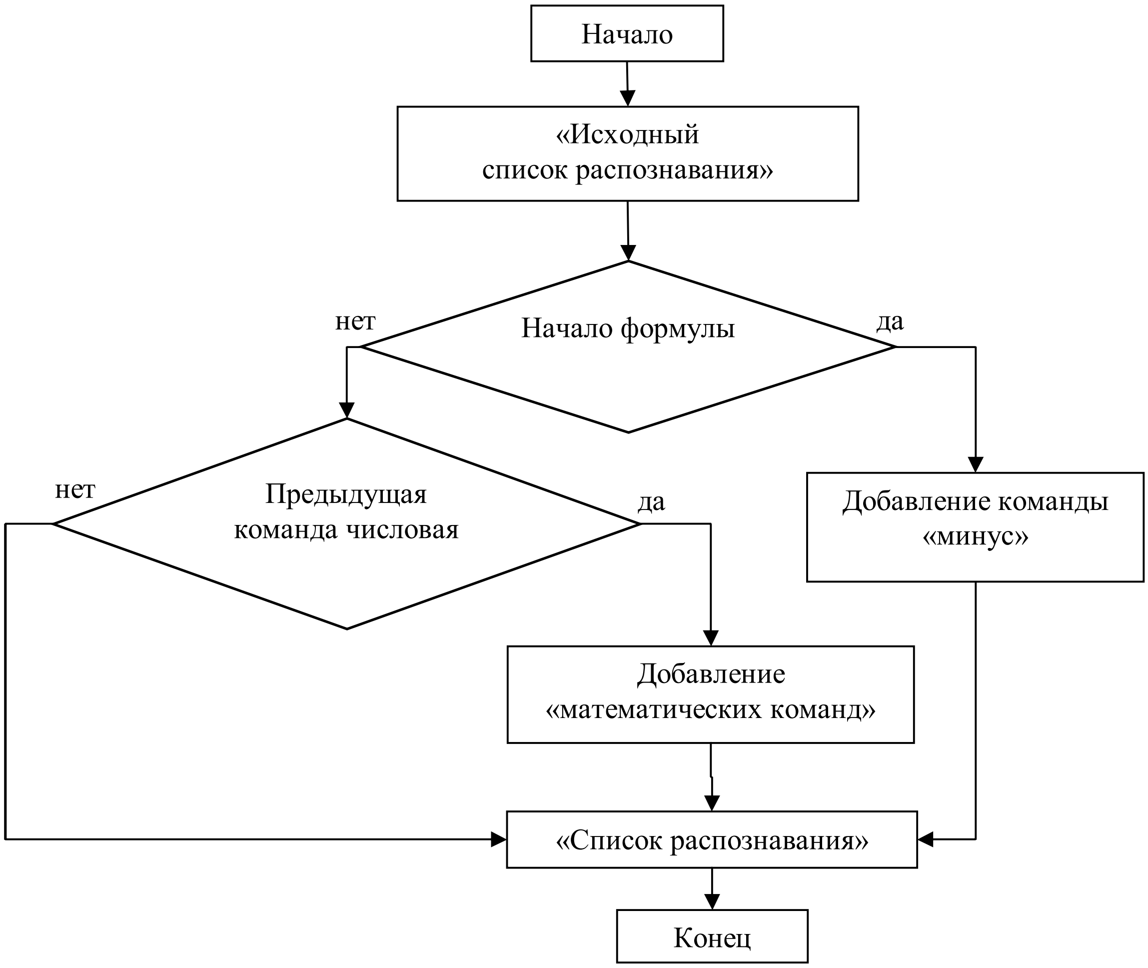


Рисунок 2 – Алгоритм формирования «списка распознавания»

Таким образом, количество команд в «списке распознавания», которые передаются в «модуль распознавания», колеблется от 10 до 20 при недоступных инженерных командах и от 28 до 40 при доступных «инженерных командах», из них в обоих случаях 5 управляющих.

Работу алгоритма формирования «списка распознавания» рассмотрим на примере ввода выражения «2+3» (табл. 2).

Таблица 2 – Пример формирования «списка распознавания».

Шаг №	Голосовая команда	Команды, добавленные к ИСР	Комментарий
1	«два»	«минус»	начало ввода формулы; к «исходному списку распознавания» добавляется команды «минус»
2	«плюс»	«математические команды»	на шаге 1 введена «числовая команда»; в «список распознавания» добавляются «математические команды»
3	«три»	–	на шаге 2 введена «математическая команда»; «список распознавания» приводится к исходному
4	«равно»	«математические команды»	на шаге 3 введена «числовая команда»; в список распознавания добавляются «математические команды»

Заключение

Задача улучшения качества распознавания в программе ГК решена за счет уменьшения количества команд в списке распознавание. Благодаря предложенному алгоритму, список распознавания формируется динамически. В результате независимого тестирования качество распознавания составило (92 ± 4) %.

Список литературы

1. Режим доступа : https://play.google.com/store/apps/details?id=sds.voice.calcrus&hl=ru
2. Шелепов В.Ю. О распознавании речи на основе межфонемных переходов / В.Ю. Шелепов, Г.В. Дорохина, А.В. Ниценко // Искусственный интеллект. –2012. – № 1 – С. 132-139.
3. Шелепов В.Ю. К проблеме распознавания слитной речи / В.Ю. Шелепов, А.В. Ниценко // Искусственный интеллект. – 2012. – № 4 – С. 272-281.
4. Шелепов В.Ю. О некоторых вопросах, связанных с дифонным распознаванием и распознаванием слитной речи / В.Ю. Шелепов, А.В. Ниценко // Искусственный интеллект. – 2013. – № 3 – С. 209-216.
5. Ниценко А.В. Сегментация и дифонное распознавание речевых сигналов / А.В. Ниценко, В.Ю. Шелепов // Системы и средства искусственного интеллекта ССИИ-2013 : материалы Международной молодежной научной школы. – 2013. – С.163-166.

References

1. https://play.google.com/store/apps/details?id=sds.voice.calcrus&hl=ru
2. Shelepov V.J, Dorokhina G.V, A.V Nicenko About speech recognition based interphoneme transitions // Artificial Intelligence . -2012 . - № 1 - p.132 -139 .
3. Shelepov V.J, Nicenko A.V On the problem of continuous speech recognition // Artificial Intelligence. - 2012 . - № 4 - p.272 -281 .
4. Shelepov V.J, Nicenko A.V On some issues related to the diphone recognition and continuous speech recognition / / Artificial Intelligence . - 2013 . - № 3 - C. 209-216.
5. Nicenko A.V, V.J Shelepov Segmentation and recognition diphone voice / / Proceedings of the International Youth Scientific School " systems and artificial intelligence SMAI 2013" . - 2013 . - P.163 - 166.

RESUME

D.S Azarenko, S.B. Ivanova, I.G. Gerasimov

Voice calculator

Background: Voice control computer is an urgent task. The "Voice Calculator" is designed for entering and calculating mathematical expressions using voice.

Materials and methods: Voice commands spoken by the user. Supporting tables for converting text commands in mathematical format.

Results: The algorithm of forming vocabulary voice commands and real-time program "Voice Calculator" which allows for voice input and calculation of simple mathematical expressions.

Conclusion: As a result, independent testing was recognition quality (92 ± 4) %.

Статья поступила в редакцию 10.04.2014.